The Positive Effects of Verbal Encouragement in Mathematics Education Using a Social Robot

LaVonda N. Brown and Ayanna M. Howard
Georgia Institute of Technology, lavonda.brown@gatech.edu, ayanna.howard@ece.gatech.edu

Abstract - **Studies have shown that the use of verbal encouragement strategies in education is able to maximize learning. This idea is derived from traditional classroom settings where teachers use a multitude of behavioral strategies to maintain the students' level of engagement. Motivated by these educational practices, we discuss the use of a Socially Interactive Robotic Tutor (SIRT) that incorporates a variety of verbal cues into multiple math learning scenarios. In this paper, we present the robotic tutor, the methods used to engage students in the learning scenario, and results from integrating the robotic tutor in the classroom environment. Results derived from 44 students engaging with SIRT during a tablet-based math test show that, when compared to non-interactive methods, verbal cues are able to increase and/or maintain student engagement regardless of student age and math content level.**

Index Terms – encouragement, math education, verbal cues, robot-based education

INTRODUCTION

In general, robotic educational agents such as a Socially Interactive Robotic Tutor (SIRT) are able to increase motivation and engagement through the use of adaptive learning techniques. Education agents are an educational method that utilizes computers or tablets as interactive teaching devices. In order for educational agents to be productive, they must do two things: monitor engagement of the student [1] and apply behavioral strategies, such as verbal or nonverbal cues, when engagement decreases [2], [3]. This is modeled from the traditional classroom setting, where teachers are able to observe a student's engagement in real-time and employ behavioral strategies to reengage the student. In effect, boredom is eliminated and attention, involvement and motivation to learn are improved [4]. This behavioral engagement is a crucial component in education because it is often related to the academic achievement of a student [5], [6].

One of the primary methods used for engagement in the classroom environment is the utilization of verbal cues, which can be used to encourage the student, provide instruction, give positive praise, etc. As such, in this paper, we focus on the impact of verbal cues in the learning environment. By changing acoustic characteristics such a tempo, pitch, intensity, voice quality, and articulation, verbal cues, or behaviors, can be used to evoke a range of emotions that impact student engagement [7]-[12]. Figure 1 summarizes how the human voice is affected by emotions such as anger, happiness, sadness, surprise, and disgust [10]. Prior research has shown that these human speech ideals can effectively be implemented on a robotic platform. For example, in [7], Breazeal was able to express emotion on a robotic platform by correlating human speech ideals (Fig. 1) into a robotic speech synthesizer. Participants were able to perceive the robot's intended emotion in most cases; however, there were a few misclassifications when the emotions shared negative valence or high arousal (i.e. angry and disgust, happy and excitement).

In addition to acoustic characteristics, sentence structure, language markers, and vocabulary choice indirectly shape the social interaction between the agent and student [8]. For example, age appropriate vocabulary is needed to maintain the student's level of engagement, and by adding markers such as "please" and "thank you," the agent can be perceived as being very polite [13]. Mutlu performed an investigation where he studied human communication and explored how robots would be able to convey the same rich social outcomes of learning, rapport, and persuasion [8]. Through combinations of verbal, vocal, and nonverbal cues, Mutlu was able to observe how embodied communication cues can be useful in enhancing social interaction in human-robot interaction (HRI).

	fear	anger	sorrow	joy	disgust	surprise
speech rate	much faster	slightly faster	slightly slower	faster or slower	very much slower	much faster
pitch average	very much higher	very much higher	slightly lower	much higher	very much lower	much higher
pitch range	much wider	much wider	slightly narrower	much wider	slightly wider	
intensity	normal	higher	lower	higher	lower	higher
voice quality	irregular voicing	breathy chest tone	resonant	breathy blaring	grumbled chest tone	
pitch changes	normal	abrupt on stressed syllable	downward inflections	smooth upward inflections	wide downward terminal inflections	rising contour
articulation	precise	tense	slurring	normal	normal	

FIGURE 1
EFFECT OF EMOTIONS ON HUMAN SPEECH [10], [11].

VERBAL BEHAVIORS IMBEDDED ON SIRT

We utilize the humanoid robot DARwIn-OP (Darwin) as the platform for our Socially Interactive Robotic Tutor

[14]. Darwin has 20 actuators, resulting in 6 DOF for each leg, 3 DOF for each arm, and 2 DOF for the neck (Fig. 2). A learning scenario includes a student and the robotic agent working together to complete a learning task, such as answering a math question. For our purposes, the learning task is achieved through a tablet-based interface, versus only using pencil and paper.

Verbal behaviors enable the educational agent to provide socially supportive phrases for reengagement as the student navigates through the learning task. During the utterance of verbal phrases, the robotic platform turns its gaze towards the student; otherwise, the robot remains looking at the tablet. The goal of the verbal phrases is to encourage the student based on their current learning performance (i.e. answering a question correctly/incorrectly; speed of answer; taking too long to answer). It is very important that the phrases are socially supportive and convey the message that the student and the robot are working together as a team. There is a dialogue established between the student and SIRT, and not a unidirectional knowledge flow (i.e. the robot is not giving instructions or issuing commands to the subject). This open dialogue integrating socially supportive phrases between teacher and student is ideal for optimal learning [15]. A sample of these socially supportive phrases is shown in Table I. For implementation purposes, the phrases were recorded using text-to-speech (TTS) software and stored on SIRT's external SD card as mp3 files.

TABLE I
SAMPLE OF VERBAL RESPONSES FROM SIRT

Answer	Speed	Phrase
Correct	Fast	"You really know your stuff!"
		"You're a genius!"
		"Fantastic!"
Correct	Slow	"This is hard, but we're doing great."
		"Thanks for all your hard work."
		"This is really making me think."
Incorrect	Fast	"Wait, I didn't get to read that one."
		"Hang in there. We're almost done."
		"I'm lost. You're going too fast."
Incorrect	Slow	"Don't worry. I had trouble with that one too."
		"That one was very challenging."
		"Don't sweat it, we'll get the next one"
None	Inactive	"Let's make an educated guess."
		"I was completely stumped on this one."
		"Don't forget about me over here."

EXPERIMENTAL DESIGN

To evaluate the effectiveness of the robotic educational agent engaging students during the learning process, we employed two between-groups design for this study. To guarantee that the skills are evenly distributed between the groups, the subjects are selected at random. A total of 24 college students took part in Trial 1 of the experiment; this consisted of both females and males in the age range of 18-33 years old (m = 24.6, σ = 4.9, Male: 18, Female: 6). A total of 20 high school students took part in Trial 2 of the experiment; this consisted of both females and males in the age range of 15-16 years old (m = 15.5, σ = 0.51, Male: 12, Female: 8). Our experiment involved one factor, type of reengagement, with two levels. Each level is defined as follows:

- **None:** This represents the control group – no agent is present.
- **Verbal:** The agent says socially supportive phrases for reengagement as the student navigates through the learning scenario. SIRT gazes towards the student when speaking to him/her; otherwise, SIRT remains looking at the tablet.

The experimental setup (Fig. 2) in this study involves a test-taking learning scenario. A Samsung Galaxy Tablet is used as the mechanism for displaying questions and recording the students' answers. The tablet is placed on an adjustable stand at eye level. For experiments with the robot agent present, SIRT is positioned to the right of the tablet, yet between the tablet and the student. The robot is placed in a position such that the robot is always able to see and interact with both the tablet and the student.

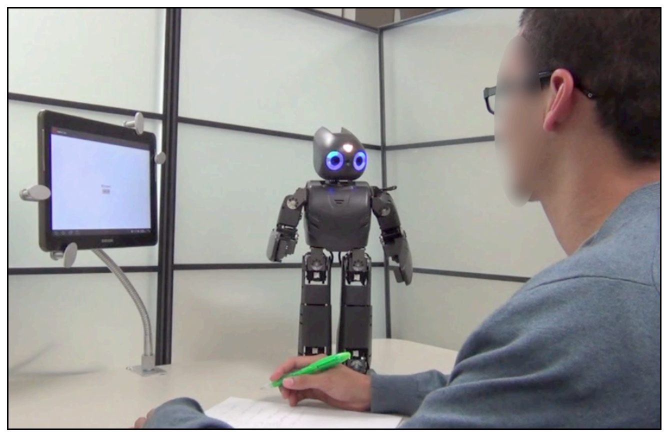

FIGURE 2
THE EXPERIMENTAL SETUP.

For experiments with the robot present, at the start of the test-taking learning scenario, SIRT gives a verbal introduction and discusses the activity that the students are about to perform. The purpose of this introduction is to eliminate the novelty of the robot from the investigation and prepare the students for the test by instructing them to gather their materials. The script of this verbal introduction is shown below:

> *"Hello. My name is Darwin. We will be going through a series of 10 math questions to learn the material together. I appreciate you taking the time out of your busy schedule to work with me. Get your pencil and paper ready so we can start. Press begin when you're ready."*

The subjects then navigate through the test questions until they reach the completion screen. The test questions consists of multiple-choice math questions, of varying difficulty levels. As each student progresses through the test, his or her interaction with the tablet is communicated to

SIRT via Bluetooth. To enable real-time performance, only the numbers 0-9 are transmitted from the tablet to the robot. Each number conveys a different message to the robot about the interaction between the student and the tablet [2]. Essentially, every button that is pressed is sent to the robot, as well as the time intervals taken to navigate through the test [2]. After test completion, a message is also sent to the robot, at which point SIRT shows its gratitude and gives a farewell.

In this study, we focus on increasing engagement while decreasing idle time by monitoring task or question duration. As such, when an answer is selected for each question (A, B, C, or D), a message is sent to the robot and it responds with the appropriate verbal behavior (for the verbal reengagement condition). An answer is classified as either being fast, slow, or average based on the time elapsed on each question: if the student submits a response in less than 30 seconds this is fast; if the student submits a response in between 30 seconds and 90 seconds this is average; if the student submits a response in more than 90 seconds this is slow. The answers are also classified based on whether or not the answer is correct. In addition, when there are long time intervals that consist of no interaction between the student and the tablet, a message is also sent to the robot. A long time interval is defined as 90 seconds; therefore, every 90 seconds of inactivity or idle time, SIRT is notified, and the robot responds appropriately.

Depending on user state, SIRT provides the users verbal cues that are selected at random based on the message sent to the robot from the tablet. For the experimental design, we utilize the same test and environmental setup across all students (so that cues happen at the same time). The only thing that changes between groups is, for the control group with no agent, the robot is simply removed from the table and not visible.

RESULTS

In this research, we look to validate the hypothesis that the use of a robotic educational agent can best increase test performance by adaptively engaging with the student using verbal encouragement. Adaptive engagement is based on the concept that the engagement model is driven by identification of the student's behavioral state. To prove or disprove this hypothesis, we first look at the different types of information that we collected separately. These include test completion time, the Likert scale questions that we ask in an exit survey, and the comments that participants left at the end of the survey.

I. Completion Time

We logged the total test time for each participant; the means for the two groups are shown in Fig. 3 & 4, and the statistical analysis is shown in Table II.

II. Survey

After the subjects completed the test, we asked them to rate their agreement with a series of statements on a 5-level Likert scale that ranged from 1 (Disagree) to 5 (Agree). One question asked for a 'yes/no/maybe' answer, which we converted to a scale from 1 (No) to 3 (Yes). For each of the questions on our survey, we performed a standard t-test to see if the changes between groups were significantly different. Table III shows the average response to each question and the p-values from the t-tests, which are separated by test groups.

III. Freeform Feedback

In addition to the questions discussed in the previous subsection, we also left room on the survey for subjects to provide freeform comments that reflected their experience as a whole. Though not everybody decided to accept the invitation, 21 of the 44 participants provided comments. Of the 11 subjects in the Verbal Category, six of them left responses about their experience. The verbal responses from both trials have been compiled into Table IV.

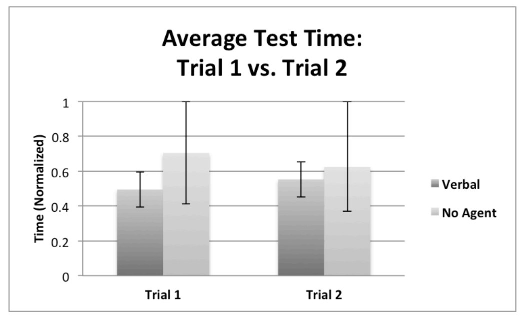

FIGURE 3
THE AVERAGE TEST COMPLETION TIMES SHOWN ALONG WITH THEIR RANGES FOR EACH TRIAL.

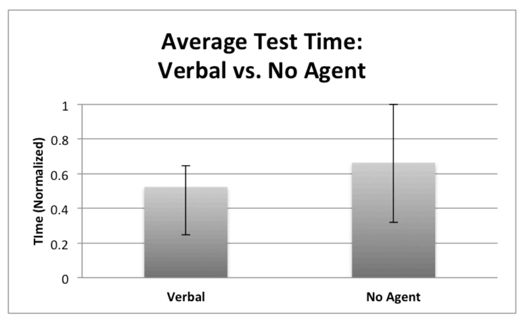

FIGURE 4
THE OVERALL AVERAGE TEST COMPLETION TIMES SHOWN ALONG WITH THEIR RANGES FOR EACH INTERACTION TYPE.

TABLE II
TOTAL TIME STATISTICAL ANALYSIS

	Mean, m	SD, σ	p-value
Trial 1:			
Verbal	0.494	0.141	0.09
No Agent	0.704	0.234	
Trial 2:			
Verbal	0.552	0.079	0.56
No Agent	0.623	0.251	
Average:			
Verbal	0.521	0.115	0.07
No Agent	0.667	0.233	

TABLE III
STATISTICAL ANALYSIS OF SURVEY RESPONSES

Question	**Trial 1:** Verbal	No Agent	p-value	**Trial 2:** Verbal	No Agent	p-value
I was nervous.	2.5	1.0	0.005*	2.2	2.0	0.83
I was frequently bored.	2.0	2.3	0.63	1.8	4.6	0.002*
I enjoyed taking the test.	3.0	4.1	0.15	4.0	2.2	0.06
I found this test difficult.	4.0	3.8	0.77	2.0	1.2	0.21
I performed better than anticipated.	3.0	3.2	0.73	2.8	2.8	1.00
I finished quickly.	3.0	2.7	0.56	3.4	2.8	0.31
This was an appropriate level for my skills.	3.3	4.3	0.12	2.6	1.8	0.39

* Statistically significant

DISCUSSION

For both trials, the verbal group was able to decrease idle time and maintain the subject's attention best with the lowest average test times. In Trial 1, when compared to the control group, the verbal group's average test time was 30% lower, while in Trial 2, the verbal group was 11% lower than the control group (Fig. 3). Across both trials, the verbal group's average test time was on average 22% lower than the control group (Fig. 4). The verbal group also presented the lowest standard deviations. In Trial 1, when compared to the control group, the verbal group's standard deviation was 40% lower, while in Trial 2, the verbal group was 69% lower than the control group. Across both trials, the verbal group's standard deviation was on average 51% lower than the control group (Fig. 4). This not only shows that the verbal cues were able to decrease time, but they were also able to do so uniformly throughout the groups. This small range, test time, and standard deviation values make it easier to guarantee a lower test completion time.

Across both trials, the survey results also show that the students in the verbal group enjoyed taking the test (m = 3.5; Slightly Agree = 4; σ = 1.51) more than the control group (m = 3.3; Neutral = 3; σ = 1.39). Also, the students were more frequently bored while taking the test in the control group (m = 3.4; Nuetral = 3; σ = 1.5) than in the verbal group (m = 1.9; Slightly Disagree = 2; σ = 1.14). Because boredom is often associated with poorer learning and behavior problems [16], it is important to note that there was a statistically significant variance in how bored the subject deemed him- or herself to be throughout the test in Trial 2. The control group with no agent present exhibited more boredom with a score of 4.6 (Agree = 5; σ = 0.55), while the verbal group was able to minimize boredom better with a score of 1.8 (Slightly Disagree = 2; σ = 1.30).

Furthermore, there was a significant variance in how nervous the subjects deemed themselves to be during the learning scenario in Trial 1. The control group was less nervous with a score of 1.0 (Disagree = 1; σ = 0), while the verbal group had an average score of 2.5 (Neutral = 3; σ = 1.05). This may be attributed to the subjects' fear of disappointing SIRT during the test as mentioned in the freeform responses in Table IV. This fear proves that the verbal cues implemented on SIRT helped to build rapport, which is needed for optimum learning [4], [17].

The freeform responses (Table IV) yield a range of responses – some students felt like the robotic platform was wasting space, while others enjoyed the robot's presence. In particular, the students said that SIRT was a "friendly looking robot with a friendly voice." Similarly, another student said SIRT was "cute…and friendly." Lastly, a student stated that he or she felt more confident when SIRT was assisting with the learning scenario. Although there were a lot of positive freeform responses, we would like make improvements in the system in the near future to decrease the amount of negative responses received from students.

CONCLUSION

In regards to minimizing idle time by actively monitoring progression through the exam, the verbal engagement implemented on SIRT was able to reach this goal best. In addition to minimizing idle time, the standard deviation was also extremely low when compared to the control group. In general, the students were less bored and enjoyed the learning scenario more when the verbal behaviors were implemented on SIRT. The students also perceived the robot to be endearing with a welcoming voice, which resulted in the students being more comfortable throughout the learning scenario.

TABLE IV
VERBAL FREEFORM RESPONSES

Trial	Response
1	"The feedback Darwin gave was really good. Even so, having someone (or something) telling you to hurry up every certain amount of time is nerve wrecking. Overall the system works great and it's very interesting."
	"I often was afraid of doing poorly on one of the questions because I didn't want Darwin to tell me that he 'found that question difficult, too.' I was mainly afraid of letting Darwin down, but I liked having him there for the test, he is a friendly looking robot with a friendly voice."
	"I was a little confused as to whether his comments meant I got the question right or wrong. He was cute though and friendly."
2	"Darwin actually made me feel more confident in myself on preforming on the test."
	"Darwin just felt like he was in the way. He was an addition that I did not need and made very little difference."
	"Darwin did provide encouragement for my efforts but he did not assist my test taking nor did he teach me any of the math skills I may have had problems on."

FUTURE WORK

This work suggests that verbal engagement is ideal for decreasing boredom and, ultimately, enhancing test performance in robot-based education (RBE). Prior studies have shown that the use of only nonverbal cues such as gestures show no significant trends when compared to verbal cues [2], [3]; therefore in the future, we would like to focus on the implementation of nonverbal behaviors to see if we can make any advancements in RBE. Lastly, we would also like to examine the effects of social engagement using robotics with children with disabilities, such as those coping with Dyscalculia or other learning disabilities.

ACKNOWLEDGMENT

This research was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903 and under Grant No. IIS-1208287. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] L. Brown and A. Howard, "A real-time model to assess student engagement for intelligent educational agents," in ASEE *submitted for publication*, 2014.

[2] L. Brown, R. Kerwin, and A. Howard, "Applying behavioral strategies for student engagement using a robotic educational agent," in IEEE SMC, 2013.

[3] L. Brown and A. Howard, "Engaging Children in Math Education using a Socially Interactive Humanoid Robot," in IEEE-RAS Humanoids, 2013.

[4] D. Szafir and B. Mutlu, "Pay attention! designing adaptive agents that monitor and improve user engagement," in CHI, May 2012.

[5] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," Review of Educational Research, vol. 74, no. 1, pp. 59–109, January 2004.

[6] C. R. Greenwood, B. T. Horton, and C. A. Utley, "Academic engagement: Current perspectives on research and practice," School Psychology Review, vol. 31, no. 3, pp. 328–349, 2002.

[7] C. Breazeal, "Emotive qualities in robot speech," in Proceedings of the 2001 International Conference on Intelligent Robotics and Systems (IROS), 2001.

[8] B. Mutlu "Designing Embodied Cues for Dialog with Robots," Association for the Advancement of Artificial Intelligence, 2012.

[9] C. Huang and B. Mutlu, "Robot Behavior Toolkit: Generating Effective Social Behaviors for Robots," in HRI, 2012.

[10] I. Murray and L. Arnott, "Toward the simulation of emotion in synthetic speech: a review of literature on human vocal emotion," Journal Acustical Society of America, vol. 93, no. 2, pp. 1097-1108, 1993.

[11] R. Picard. Affective Computation, MIT Press, Cambridge, MA, 1997.

[12] J. Cahn, "Generating Expression in synthesized speech," M.S. thesis, MIT Media Lab, Cambridge, MA, 1990.

[13] P. Brown and S. Levinson, "Politeness," Cambridge Univ. P., 1998.

[14] I. Ha, Y. Tamura, H. Asama, J. Han, and D. Hong, "Development of open humanoid platform darwin-op," in SICE, 2011, pp. 2178–2181.

[15] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse, "Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor," in CHI, April 2010, pp. 1613–1622.

[16] Baker, R. S. J., D'Mello, S. K., Rodrigo, M. M. T., Graesser, A. C.: Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments, International Journal of Human-Computer Studies, pp. 223–241 (2010)

[17] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in IROS, September 2008, pp. 3230–3235.

AUTHOR INFORMATION

LaVonda Brown, Ph.D. student, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

Ayanna M. Howard, Motorola Foundation Professor, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.